2008 International Symposium on Nonlinear Theory and its Applications
NOLTA'08, Budapest, Hungary, September 7-10, 2008

Potpourri of Applications of Complex Networks Research

Chi K. Tse

Department of Electronic and Information Engineering
The Hong Kong Polytechnic University, Hong Kong
URL: http://chaos.eie.polyu.edu.hk

Abstract— Research on complex networks has been a subject of rigorous theoretical research in the mathematics and physics research communities in the past decade. The many discoveries that man-made and natural networks display a power-law degree distribution and small-world property have clearly indicated a high level of relevance of the study of complex networks with real-world applications. In this paper some recent results in applying complex networks research in real-world problems will be reviewed. The emphasis is on how complex networks would provide a new perspective on the way problems can be formulated, leading to possible new solution approaches. Examples in engineering, disease transmission, language, music and finance are given.

1. Introduction

In the past decade, complex networks have attracted a great deal of attention from researchers across a variety of disciplines including mathematics, science, engineering and humanities. Certain classes of complex problems, arising from many different disciplines, have been analyzed from a networking viewpoint. Results generated from such network-based analyses often yield new insights into the basic structure of the system under study as well as the way in which the various subsystems interact. The basic foundation of the analysis is that the system can be broken down into a large number of basic units or subsystems which are interconnected with one another, and specifically, the way in which connections are distributed over the entire system plays an important role in determining the behavior of the whole system. The basic units or subsystems can be identical or different. Although a considerable amount of fundamental findings have been reported in complex networks, such as the general scalefree and small-world properties of networks arising from human interactions, man-made and natural networks, the progress of applying complex network analysis to practical problems is still relatively slow. In this paper we briefly review a few applications. Our purpose is to show some possible pathways through which practical problems may be tackled from a network viewpoint, yielding entirely new insights into the problems. We will first present a brief overview of networks, and present a few cases of problem formulation in terms of networks, including telephone traffic analysis, disease transmission dynamics, music composition and stock market fluctuation.

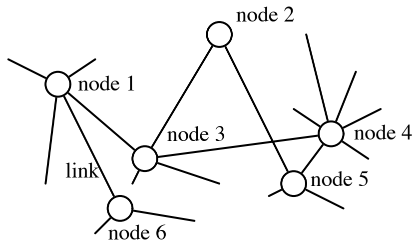

Figure 1: Network.

2. Overview of Networks

A network is a collection of "nodes" connected by "links" or "edges" [1, 2]. An illustration is shown in Fig. 1. For the purpose of this paper, we do not impose any restriction on the way a node should be defined. Basically, a node can be any basic element that can be identified in a system which is being considered as a network. Such basic nodes are usually very numerous in quantity and are inter-connected via physical couplings or abstract connections. A simple example is a human community, where people can be regarded as nodes. If two persons know each other, we put a link connecting these two persons. Then, a network can be constructed for this community. In this network, nodes are people, and links are some sorts of relationship. The resulting network topology will depend on how we actually define the links. As another example, if we are studying the transmission of avian influenza among the birds in different cities or locations, we may consider a city or place as a node where outbreak of the disease occurs. We may then define a criterion for the purpose of connecting any two nodes. For instance, we put a link between two cities if outbreaks occur in these two cities within a certain distance in both space and time. Then, the resulting network represents a kind of "map" of transmission of the disease. Finally, consider a piece of music. Suppose we wish to construct a network for a given piece of music. We may define a node as a musical note and the way in which individual musical notes are connected together defines a particular piece of work. One straightforward way to define a link is co-occurrence, i.e., two musical notes are connected if they appear adjacent to each other. Thus, a music score can be represented by a network and we may study music in terms of the property of this network. We may further conceive that any complex system can be considered as a complex network by defining nodes and links appropriately.

The number of links emerging from and converging at a node is called the "degree" of that node, usually denoted by k. So, we have an average degree for the whole network. The key concept here is the distribution of k. This concept can be mathematically presented in terms of probability density function. Basically, the probability of a node having a degree k is $p(k)$, and if we plot $p(k)$ against k, we get a distribution function. This distribution tells us about how this network of nodes are connected. For instance, if we see a normal distribution of $p(k)$ with mean k_m, that means most nodes are having an average degree k_m, some having less and some having more, following a Gaussian type of spread about k_m. Recent research has provided concrete evidence that networks with man-made couplings and/or human connections follow power-law distributions, i.e., $p(k)$ vs k being a straight line whose gradient is the characteristic exponent normally denoted by γ [3]–[8]. Such networks are termed *scalefree networks*.

3. Examples of Applications

3.1. Telephone Traffic Analysis [10]–[12]

Traffic analysis is an important topic in communication engineering. Specifically, the usual problem of traffic analysis of a telephone network is to analyze the extent to which a given telephone can support the communication demand in terms of the amount of calls that can be carried by the network per a defined time duration. An important part of the analysis is in modeling the way in which calls are made among users. Calculation of the amount of call failures or blockings is an essential part of a traffic analysis. Conventional models basically do not consider the properties of the user community and simply assume a regular or democratic user relationship, i.e., all users are regularly or evenly connected and may make calls with one another with an assumed identical probability. Then, with an assumed set of call characteristics, such as call duration distribution, channel capacity, etc., analysis can be carried out to evaluate the carried traffic.

Here, recent results in complex networks are directly applicable in refining (actually with quite drastic improvement) the aforementioned traffic analysis. Specifically, the users of the telephone network are a network of people! It makes no sense to assume that users are regularly or evenly connected, and in reality a user normally only calls his/her own acquaintances, such as friends, family members, relatives, etc. Thus, we may make use of the user network property to improve results of traffic analysis [10, 11]. Precisely, we construct a user network, which has been found to follow a scalefree degree distribution [12], and use this network for traffic analysis. Fig. 2 shows the degree distribution of a user network which has been constructed for this purpose. Results from such analyses have shown that,

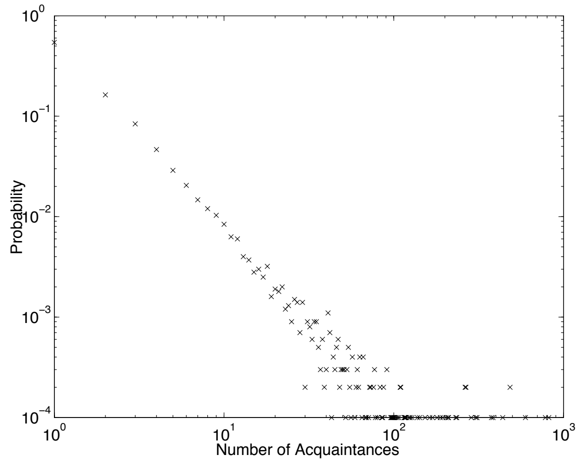

Figure 2: Degree distribution of user network. This plot basically shows the probability of a user having n acquaintances versus n. The average n is 5.

1. the carried traffic is drastically reduced as a result of the presence of a small number of very heavy users or supernodes when a scalefree user degree distribution is assumed (see Fig. 3) [10];
2. increase in channel capacity helps improve the carried traffic only up to a certain threshold value, beyond which the carried traffic cannot be improved. When the user network is scalefree, the threshold becomes significantly lower (see Fig. 4) [11].

The above results clearly pinpoint the impact of user behavior on the traffic of a telephone network. Call failures are generally more severe in the real world when users are scalefree distributed, and increasing channel capacity (investment) may not be the best solution to alleviating traffic. Other pricing strategies could be more effective.

3.2. Disease Transmission

Conventional study of geographical disease transmission assumes that the transmission paths are locally homogeneous and that the pathogen will diffuse uniformly. Typical mathematical formulation involves a standard diffusion equation that describes the dynamics of the transmission [13]. A natural consequence of this basic assumption is that the disease will terminate if the pathogen transmits at a rate lower than a threshold. Recent studies have showed that the topology of the transmission plays an important role, and in particular if the transmission network follows a scalefree distribution and the pathogen transmits fast enough, then an epidemic is inevitable. In [14], we consider a network of cities or locations where an outbreak of avian influenza was reported between 25 November 2003 and 10 March 2007. The nodes are the different cities or locations of outbreaks,

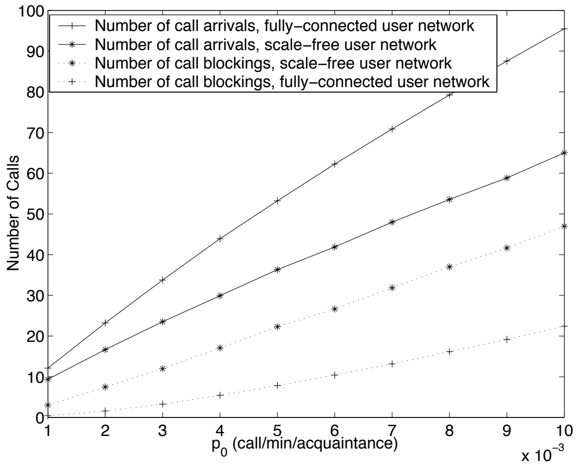

Figure 3: Call arrivals versus call rates, comparing fully-connected user network with scalefree user network. The call arrivals are reduced significantly for a scalefree user network. p_0 is the call rate [10].

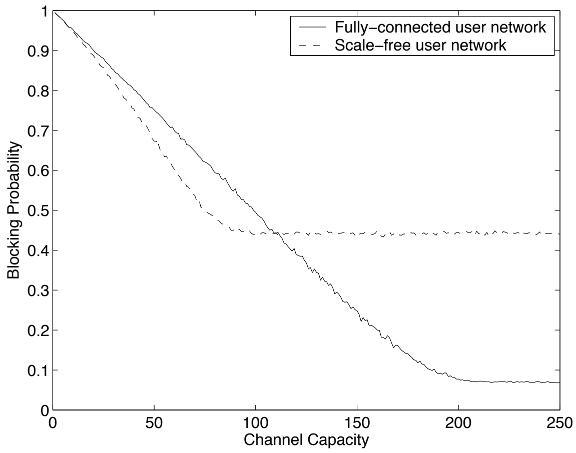

Figure 4: Call blocking (failure) probability versus channel capacity. This shows that channel capacity is no longer the determining factor when it has reached a threshold value. The case of scalefree user networks has a much lower threshold, showing that this type of networks has a more severe blocking problem [11].

and two nodes are connected by possible pathways of transmission of the virus. A very simple connection criterion has been used, which basically puts a link between two nodes if the geographical distance between the two nodes is less than the distance traversed by the virus at an assumed rate μ (km per day) within the time difference between the occurrence of the outbreaks at the two nodes, i.e.,

$$d(i,j) < (t_i - t_j)\mu$$

where $d(i,j)$ is the distance between node i and node j, and t_i, t_j are the times of outbreaks and $t_i > t_j$. The distance $d(i,j)$ is actually calculated from the actual latitudes and longitudes of nodes i and j [14]. It has been found that the degree distribution of the network is scalefree.

The scalefree transmission topology has a clear implication on the transmission dynamics; that is, eradication of that disease is only possible if transmission is reduced to precisely zero. Other studies related to applications of complex networks in disease transmission have also produced insightful results which are not provided by conventional studies [15, 16].

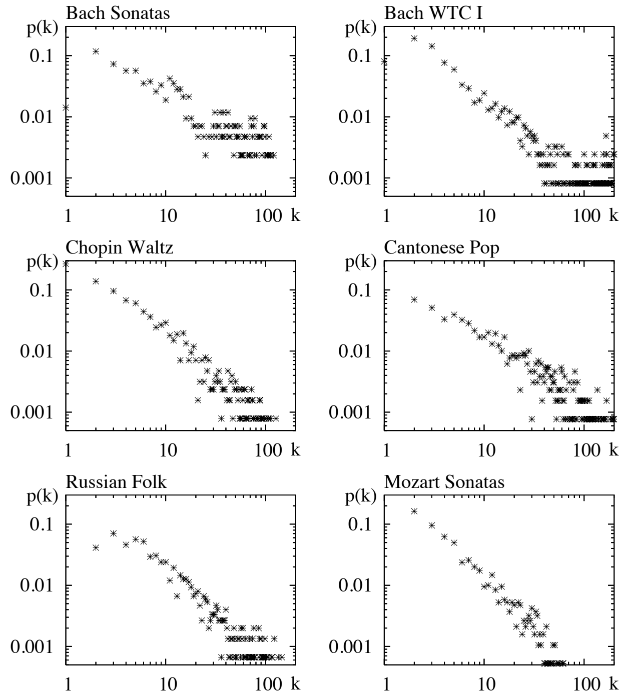

Figure 5: Degree distributions for different music. $p(k)$ versus k in log-log scale.

3.3. *Music*

The network approach can also be applied to analyze music. A recent attempt is to construct networks for music based on co-occurrence of musical notes. It has been found that regardless of the types of music, networks formed by connecting co-occurring musical notes display a scalefree degree distribution [17]. Such scalefreeness can thus be regarded as a basic property of music that is appealing to the human perception. Based on this property we may construct music artificially. Some samples of music generated from the musical networks can be downloaded from the the following website:

- http://cktse.eie.polyu.edu.hk/MUSIC/

3.4. *Stock Market*

The structure of stock markets can be effectively examined from the complex network viewpoint. In [18], we consider 19807 US stocks (all the US stocks that were traded

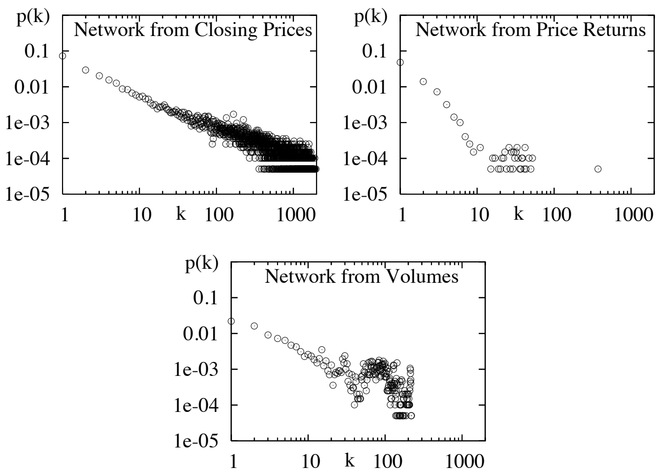

Figure 6: Scalefree degree distribution of networks formed by connecting stocks based on cross correlation of daily closing prices (upper left panel); daily price returns (upper right panel); and daily trading volumes (lower panel).

each trading day from July 1, 2005 to August 30, 3007). The construction procedure is based on connecting any two stocks whose daily price time series are "similar" in terms of cross correlation evaluated over a period of time. For the first time, full network data of all US stocks traded each trading day over a 2-year period have been reported in [18, 19]. It has been found that the networks formed using high cross correlation as the connection criterion are scalefree. Fig. 6 shows the degree distribution of networks formed by connecting stocks based on cross correlation of daily closing prices, daily price return, as well as daily trading volumes. The results suggest that a relatively small number of stocks are exerting much of the influence over the majority of stocks. In [19], we have also shown that stock market fluctuation may be detected from the variation of the degree distribution of the network formed from the correlation network.

4. Conclusion

A great number of possibilities exist for applying complex network modeling to study complex systems. Due to space limitation in this paper, we have presented a few examples to illustrate some possible applications. In brief, the definition of nodes and the choice of connection criteria determine the network construction which in turn determines how the resulting network may be used in the analysis of the corresponding complex system.

Acknowledgment: The works described in this paper were supported by Hong Kong Polytechnic University grant 1-BBZA.

References

[1] A.-L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science,* vol. 286, pp. 509–512, Oct. 1999.

[2] S. H. Strogatz, "Exploring complex networks," *Nature,* vol. 410, pp. 268–276, March 2001.

[3] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Aberg, "The web of human sexual contacts," *Nature,* vol. 411, pp. 907–908, June 2001.

[4] M. E. J. Newman, "Scientific collaboration networks I: Network construction and fundamental results," *Physical Review E,* vol. 64, pp. 016131-1-8, 2001.

[5] M. E. J. Newman, "Scientific collaboration networks II: Shortest paths, weighted networks, and centrality," *Physical Review E,* vol. 64, pp. 016132-1-7, 2001.

[6] G. Csanyi and B. Szendroi, "Structure of a large social network," *Physical Review E,* vol. 69, pp. 036131-1-5, 2004.

[7] S. Battiston and M. Catanzaro, "Statistical properties of corporate board and director networks," *European Physical Journal B,* vol. 38, pp. 345-352, 2004.

[8] G. Ravid and S. Rafaeli, "Asynchronous discussion groups as small world and scale free networks," *Peer-Reviewed Journal on the Internet,* vol. 9, no. 9, 2004.

[9] X. Zheng, F. C. M. Lau, and C. K. Tse, "Study of LPDC codes built from scale-free networks," *Proc. Int. Symp. Nonlinear Theory and Its Applications,* Bologna, Italy, pp. 563–566, September 2006.

[10] Y. Xia, C. K. Tse, W. M. Tam, F. C. M. Lau, and M. Small, "Scale-free user-network approach to telephone traffic analysis," *Physical Review E,* vol. 72, 026116-1-7, August 2005.

[11] Y. Xia, C. K. Tse, W. M. Tam, F. C. M. Lau, and M. Small, "Analysis of telephone network traffic based on a complex user network," *Physica A,* vol. 368, Issue 2, pp. 583–594, August 2006.

[12] W. M. Tam, F. C. M. Lau, and C. K. Tse, "Complex-network modeling of a call network," *IEEE Transactions on Circuits and Systems I,* to appear.

[13] J. D. Murray, *Mathematical Biology,* New York: Springer, 1993.

[14] M. Small, D. M. Walker, and C.K. Tse, "Scale free distribution of avian influenza outbreaks," *Physical Review Letters,* vol. 99, 188702-1–4, November 2007.

[15] M. Small, C. K. Tse, and D. M. Walker, "Super-spreaders and the rate of transmission of the SARS virus," *Physica D,* vol. 215, pp. 146–158, March 2006.

[16] M. Small and C. K. Tse, "Small world and scale free network model of transmission of SARS," *International Journal of Bifurcation and Chaos,* vol. 15, no. 5, pp. 1745–1756, May 2005.

[17] C. K. Tse, X. Liu, and M. Small, "Analyzing and composing music with complex networks: finding structures in Bach's, Chopin's and Mozart's," *International Symposium on Nonlinear Theory and Its Applications,* (NOLTA2008), Budapest, Hungary, September 2008.

[18] C. K. Tse, J. Liu, and F. C. M. Lau, "Winner-take-all correlation-based complex networks for modeling stock market and degree-based indexes," *International Symposium on Nonlinear Theory and Its Applications,* (NOLTA2008), Budapest, Hungary, September 2008.

[19] J. Liu, C. K. Tse, and K. He, "Detecting stock market fluctuation from stock network structure variation," *International Symposium on Nonlinear Theory and Its Applications,* (NOLTA2008), Budapest, Hungary, September 2008.